ENTERED

OIPE

RAW SEQUENCE LISTING   DATE: 06/03/2002
PATENT APPLICATION: US/10/045,180A   TIME: 12:40:19

Input Set : A:\EP.txt
Output Set: N:\CRF3\06032002\J045180A.raw

```
 3 <110> APPLICANT: Bougueleret, Lydie
 4         Chumakov, Ilya
 6 <120> TITLE OF INVENTION: Human Defensin Polypeptide Def-X, Genomic DNA and cDNA,
Composition
 7         Containing Them and Applications to Diagnosis and to Therapeutic Treatment
 9 <130> FILE REFERENCE: GEN-100D1
11 <140> CURRENT APPLICATION NUMBER: US 10/045,180A
12 <141> CURRENT FILING DATE: 2001-10-18
14 <150> PRIOR APPLICATION NUMBER: US 09/486,580
15 <151> PRIOR FILING DATE: 2000-02-25
17 <150> PRIOR APPLICATION NUMBER: PCT/FR98/01864
18 <151> PRIOR FILING DATE: 1998-08-28
20 <150> PRIOR APPLICATION NUMBER: FR 97/10823
21 <151> PRIOR FILING DATE: 1997-08-29
23 <160> NUMBER OF SEQ ID NOS: 12
25 <170> SOFTWARE: PatentIn version 3.1
27 <210> SEQ ID NO: 1
28 <211> LENGTH: 4415
29 <212> TYPE: DNA
30 <213> ORGANISM: Homo sapiens
32 <220> FEATURE:
33 <221> NAME/KEY: misc_feature
34 <222> LOCATION: (1)..(4415)
35 <223> OTHER INFORMATION: Def-X genomic sequence
38 <220> FEATURE:
39 <221> NAME/KEY: misc_feature
40 <222> LOCATION: (85)..(85)
41 <223> OTHER INFORMATION: n = a, c, g, or t.
44 <220> FEATURE:
45 <221> NAME/KEY: misc_feature
46 <222> LOCATION: (143)..(143)
47 <223> OTHER INFORMATION: n = a, c, g, or t.
50 <220> FEATURE:
51 <221> NAME/KEY: misc_feature
52 <222> LOCATION: (670)..(670)
53 <223> OTHER INFORMATION: n = a, c, g, or t.
56 <220> FEATURE:
57 <221> NAME/KEY: misc_feature
58 <222> LOCATION: (970)..(970)
59 <223> OTHER INFORMATION: n = a, c, g, or t.
62 <220> FEATURE:
63 <221> NAME/KEY: misc_feature
64 <222> LOCATION: (1111)..(1111)
65 <223> OTHER INFORMATION: n = a, c, g, or t.
```

RAW SEQUENCE LISTING   DATE: 06/03/2002
PATENT APPLICATION: US/10/045,180A   TIME: 12:40:19

Input Set : A:\EP.txt
Output Set: N:\CRF3\06032002\J045180A.raw

```
 68 <220> FEATURE:
 69 <221> NAME/KEY: misc_feature
 70 <222> LOCATION: (1150)..(1150)
 71 <223> OTHER INFORMATION: n = a, c, g, or t.
 74 <220> FEATURE:
 75 <221> NAME/KEY: CAAT_signal
 76 <222> LOCATION: (1711)..(1714)
 77 <223> OTHER INFORMATION:
 80 <220> FEATURE:
 81 <221> NAME/KEY: TATA_signal
 82 <222> LOCATION: (1758)..(1767)
 83 <223> OTHER INFORMATION:
 86 <220> FEATURE:
 87 <221> NAME/KEY: misc_feature
 88 <222> LOCATION: (1780)..(1780)
 89 <223> OTHER INFORMATION: n = a, c, g, or t.
 92 <220> FEATURE:
 93 <221> NAME/KEY: misc_feature
 94 <222> LOCATION: (1836)..(1874)
 95 <223> OTHER INFORMATION: Exon 1
 98 <220> FEATURE:
 99 <221> NAME/KEY: misc_feature
100 <222> LOCATION: (1875)..(1880)
101 <223> OTHER INFORMATION: splice donor site
104 <220> FEATURE:
105 <221> NAME/KEY: misc_feature
106 <222> LOCATION: (1974)..(1974)
107 <223> OTHER INFORMATION: n = a, c, g, or t.
110 <220> FEATURE:
111 <221> NAME/KEY: misc_feature
112 <222> LOCATION: (2117)..(2117)
113 <223> OTHER INFORMATION: n = a, c, g, or t.
116 <220> FEATURE:
117 <221> NAME/KEY: misc_feature
118 <222> LOCATION: (2133)..(2133)
119 <223> OTHER INFORMATION: n = a, c, g, or t.
122 <220> FEATURE:
123 <221> NAME/KEY: misc_feature
124 <222> LOCATION: (2155)..(2335)
125 <223> OTHER INFORMATION: Alu insertion
128 <220> FEATURE:
129 <221> NAME/KEY: misc_feature
130 <222> LOCATION: (2186)..(2186)
131 <223> OTHER INFORMATION: n = a, c, g, or t.
134 <220> FEATURE:
135 <221> NAME/KEY: misc_feature
136 <222> LOCATION: (2191)..(2191)
137 <223> OTHER INFORMATION: n = a, c, g, or t.
140 <220> FEATURE:
```

RAW SEQUENCE LISTING   DATE: 06/03/2002
PATENT APPLICATION: US/10/045,180A   TIME: 12:40:19

            Input Set : A:\EP.txt
            Output Set: N:\CRF3\06032002\J045180A.raw

141 <221> NAME/KEY: misc_feature
142 <222> LOCATION: (2367)..(2367)
143 <223> OTHER INFORMATION: n = a, c, g, or t.
146 <220> FEATURE:
147 <221> NAME/KEY: misc_feature
148 <222> LOCATION: (2710)..(2780)
149 <223> OTHER INFORMATION: L1 fragment insertion
152 <220> FEATURE:
153 <221> NAME/KEY: misc_feature
154 <222> LOCATION: (3391)..(3393)
155 <223> OTHER INFORMATION: splice acceptor site
158 <220> FEATURE:
159 <221> NAME/KEY: misc_feature
160 <222> LOCATION: (3394)..(3577)
161 <223> OTHER INFORMATION: Exon 2
164 <220> FEATURE:
165 <221> NAME/KEY: misc_feature
166 <222> LOCATION: (3406)..(3408)
167 <223> OTHER INFORMATION: Translation initiation codon (ATG)
170 <220> FEATURE:
171 <221> NAME/KEY: misc_feature
172 <222> LOCATION: (3578)..(3583)
173 <223> OTHER INFORMATION: splice donor site
176 <220> FEATURE:
177 <221> NAME/KEY: misc_feature
178 <222> LOCATION: (4123)..(4123)
179 <223> OTHER INFORMATION: n = a, c, g, or t.
182 <220> FEATURE:
183 <221> NAME/KEY: misc_feature
184 <222> LOCATION: (4161)..(4163)
185 <223> OTHER INFORMATION: splice acceptor site
188 <220> FEATURE:
189 <221> NAME/KEY: misc_feature
190 <222> LOCATION: (4164)..(4379)
191 <223> OTHER INFORMATION: Exon 3
194 <220> FEATURE:
195 <221> NAME/KEY: misc_feature
196 <222> LOCATION: (4274)..(4276)
197 <223> OTHER INFORMATION: Translation termination codon (TAA)
200 <220> FEATURE:
201 <221> NAME/KEY: polyA_signal
202 <222> LOCATION: (4374)..(4379)
203 <223> OTHER INFORMATION:
206 <400> SEQUENCE: 1
207 acaccatttg tcttcatgta accccattag ctataccctc tagtgcaagg aaaccatagg   60
W--> 209 gcctaggtca caccatgagg ctgcncttac aagttatgca aaaactatgg acttgggaga   120
W--> 211 cctgtgcgta acaacatcac acnccaaatt taaccagctc tccccataac agcacgctca   180
213 tgtgttactg aggaaatgcc tgtggattgg agtgtgttct gtgtgcagga ggctggtcca   240
215 ggtttcactt ctgcaggaca ctggacgttt cccaaaacca gcagactttc cccacgtgca   300

RAW SEQUENCE LISTING   DATE: 06/03/2002
PATENT APPLICATION: US/10/045,180A   TIME: 12:40:19

Input Set : A:\EP.txt
Output Set: N:\CRF3\06032002\J045180A.raw

```
       217 cacacacccc ttctcatttt gcctctacat ccatatccac tgggcccttc aggcacctac    360
       219 taatgcccta gaacctaaaa ccatcatctg gggcccagtt ccctgaatgg ccctaatctc    420
       221 ttcctctgct ggaatgagtc cagtgcccac ttcctccaac ggtgaaattg ctgggctgct    480
       223 acagatcagg aactcactgc ttcctcatag gggcagccga cttcactgct ctgcaacagc    540
       225 gaccacccct agcgaggctt gagatgcctc ttgcctcctt aagactgagg gagacgcttc    600
       227 agctctcact ccactgcccc aagtcctcca cagcgcggtg cctgctgcct tcacacagag    660
W--> 229 ctgcaggggn aggtcctgtg tatccggcct gctggaccag cgctgtgcac aaccctccca    720
       231 tggcaacagt ggctgcccgg cctgcacact gggcttggca acctcgctgt aggtatttat    780
       233 tccctcagga gtgactgcat tcttttccca tttccagaaa actgatgcca tttacctcac    840
       235 tatgaggagg aggaggagga ggagggtgga gagtggtaca ttttaaaatg tgcactattc    900
       237 tccctaggac tccccctcaa ataacccagg agggaccata ccagctcatt cctgtgtatc    960
W--> 239 ccaagcatan gagtaatcat cccactcatg ctgagtgtat ggtggccatt aagcctgccc   1020
       241 tgaactggct ttagaacaag gtgtttgagc acacagcacc gtcttgctgc caccttggcc   1080
W--> 243 ccctcccttg tgagacctct gagacacatt naggtctcac ctaaaaatct caggatttct   1140
W--> 245 aggcccaaan cggtcctaaa aaattgttca gtctgaactc tctaaggtca agagaagagg   1200
       247 tggttgctcc ctctaagaaa ccacatgttg catgtacatc cttaattccg gaaagtccaa   1260
       249 caaacctgcc ctgcttagca acacaagccg aggtggtact cctctcaccc gggcattctc   1320
       251 caacacacct gtttgtccaa acagctttga tttgttttta tagttggacc ccaggttccc   1380
       253 aggaggctgg ttcaggccat attccaaatc ctcatctgtg tgtgagtggc attcttagcc   1440
       255 tagcctcctt acagggtgga tactatgata cacagccagg ctgtcccagt ggctttcaat   1500
       257 attcttttgg tccagatagt tcagcctcag caccagtgta ggcatcacag ggtcaattgt   1560
       259 cttaggagtc atggagaatt catagttggt agctacctgg gcctggccag ggctgaccat   1620
       261 agacaaggca tccctctgtg aactcctatt ttaatgccag cttcccaaca aatttctcaa   1680
       263 ctgctcttac cagcaggtat ttaaactact caatagaaag taaccctgaa aattaggaca   1740
W--> 265 cctgttccca aaagaccctt aaatagggga agtcctttcn ctgcttgtgc acagctgctg   1800
       267 atgtggcaac atgaggcctg ggacagggga ctgtcctctg cccactctgg tagcctcacg   1860
       269 tagcttaaca atctgtcagt aatacaatac aaaacttaaa ctttcatact gcggttccac   1920
W--> 271 ccaggaagct gtgttcccaa tctgacccgt gattatgggg ccacctcaga gggnacccag   1980
       273 tgagggaata ttttgccatc tgggactgtt ggttgctggg ggcagtggct atgagctcag   2040
       275 ttaataaact caagcagttt ccttccaaac acacatgtcc tacttaacgt gtccaacaga   2100
W--> 277 gatgatcata ctcatangct gctaaaacat tantttttt ttgagaaaag tctattcatg   2160
W--> 279 ttcttggccc atggagtttt catttnatta ntttatttat tttgcagaga tggagtctca   2220
       281 ctatgttgct caagctggtc tccaactcct gggctcaagc gatcttccta ctttggcctt   2280
       283 tgaaagcgct gagattgcct gtgtgagcca tcatgggggc tcactggccc actgattaat   2340
W--> 285 cagattaatt gttttttgct attgaanttg tttgacttcc ttgtatattc ggatatttac   2400
       287 ccattctaac acgtagggtt tgcaaatatt ttctctcatg ttctgtgttg ccttttcact   2460
       289 cagttgatgg tttcctttgc tgtgcaggtg ctttagtgtt caacgcagcc ccgcttgtct   2520
       291 attttccatt ttattgcctg tccctttgat gtcatagcca agaaataatt gcccagatta   2580
       293 atgtcaaaaa gctttatccc tatatattct tctagtagtt tatggtttca gatcttatgt   2640
       295 ttaggtcttc aatccattga gttgattttt gtatgtggta taagaaaaaa gaccacatgt   2700
       297 atacatatct caaattctaa ggtagtatat attagacaca tacaatgtgt ctatttacac   2760
       299 acattgagct gaaaataata aacatatttt tatctttcaa tcaactctat ctctatctca   2820
       301 ctgaacttgt ttcacctata gcctgatgag gttgctgtcc tctctacccc agctcctata   2880
       303 ggagactgct catcccctaa cctcaaaaac cccttcatga gggtgataat gcccttgaat   2940
       305 cctgcaatga attagttctc tactacagtg gaattcaggt ctgttatgag ggtctggatc   3000
       307 tctgaagaga agagctctca tttttcagaaa ataagcagga tttattccct gaaattactg   3060
       309 aattaaatca ctgtttcgat tacttttttgc aatattaaaa gtaaatattt aaacaggtaa   3120
       311 aaacagaaat aatggtaggg tccttatcat caccgtgaat tccaagctag catagacact   3180
       313 aaacctagag attcacacta gaatgaaagc tgggagagca gaggagtctc agaaggatgt   3240
```

```
          RAW SEQUENCE LISTING              DATE: 06/03/2002
          PATENT APPLICATION:  US/10/045,180A    TIME: 12:40:19

              Input Set : A:\EP.txt
              Output Set: N:\CRF3\06032002\J045180A.raw
```

```
      315 ggaggccaat ggacacctgc aacctctcca acgaaatgcc tacctcctct cactgcagca   3300
      317 tccatctctg agccttctcg cagcagagct ataaattcag cctggctcct ccgttcccac   3360
      319 acatccactc ctgctctccc tcctctcctc caggtgacta cagttatgag gaccctcacc   3420
      321 ctcctctctg cctttctcct ggtggccctt caggcctggg cagagccgct ccaggcaaga   3480
      323 gctcatgaga tgccagccca gaagcagcct ccagcagatg accaggatgt ggtcatttac   3540
      325 ttttcaggag atgacagctg ctctcttcag gttccaggtg agagatgcca gcatgcagag   3600
      327 ctacagacta gacagaagga caggagacag gctctggaat tggatctcag tggcagatgt   3660
      329 cacttaggtg gctatactta acatctctgg tcctggattt tctcatatct aaatggaata   3720
      331 gagaaccaaa gaaatctaag agatttttct ttctccaaaa acttgattcc aagatatgac   3780
      333 tgtgaaattc actagattta agatataagg agatgctacc tagttccttc tggagccaga   3840
      335 caaacaagct taagtatata ggaaaatatt tcaccctgtc tatataggag gttttagaac   3900
      337 ctggagagga gcctaagaat gtgttcaggt gtgtgtgtga tgggcaggaa tgcagaaaag   3960
      339 tgaagcaaag gagaatgagt ctcgaatcct gtgtgaccag cactgctctg tgtatttatt   4020
      341 cctattgact gagattgttt gtgctaccgg ctgtaataca gccaacatca ctcatcagcc   4080
W-->  343 aacatgtgac ttctccaaga ttcccttac cacccactgc tgnaccccgt actcagtttc   4140
      345 tgatgctctc tctgggtccc caggctcaac aaagggcttg atctgccatt gcagagtact   4200
      347 atactgcatt tttggagaac atcttggtgg gacctgcttc atccttggtg aacgctaccc   4260
      349 aatctgctgc tactaagctt gcagactaga gaaaaagagt tcataatttt ctttgagcat   4320
      351 taaagggaat tgttattctt ataccttgtc ctcgatttcc tgtcctcatc ccaaataaat   4380
      353 acttggtaac atgatttccg ggtttttttt ttttt                               4415
      356 <210> SEQ ID NO: 2
      357 <211> LENGTH: 453
      358 <212> TYPE: DNA
      359 <213> ORGANISM: Homo sapiens
      361 <220> FEATURE:
      362 <221> NAME/KEY: CDS
      363 <222> LOCATION: (52)..(336)
      364 <223> OTHER INFORMATION: Def-X coding sequence
      367 <400> SEQUENCE: 2
      368 ctctgcccac tctggtagcc tcacgtagct taacaatctg tgactacagt t atg agg    57
      369                                                           Met Arg
      370                                                            1
      372 acc ctc acc ctc ctc tct gcc ttt ctc ctg gtg gcc ctt cag gcc tgg   105
      373 Thr Leu Thr Leu Leu Ser Ala Phe Leu Leu Val Ala Leu Gln Ala Trp
      374         5                   10                  15
      376 gca gag ccg ctc cag gca aga gct cat gag atg cca gcc cag aag cag   153
      377 Ala Glu Pro Leu Gln Ala Arg Ala His Glu Met Pro Ala Gln Lys Gln
      378         20                  25                  30
      380 cct cca gca gat gac cag gat gtg gtc att tac ttt tca gga gat gac   201
      381 Pro Pro Ala Asp Asp Gln Asp Val Val Ile Tyr Phe Ser Gly Asp Asp
      382 35                  40                  45                  50
      384 agc tgc tct ctt cag gtt cca ggc tca aca aag ggc ttg atc tgc cat   249
      385 Ser Cys Ser Leu Gln Val Pro Gly Ser Thr Lys Gly Leu Ile Cys His
      386                 55                  60                  65
      388 tgc aga gta cta tac tgc att ttt gga gaa cat ctt ggt ggg acc tgc   297
      389 Cys Arg Val Leu Tyr Cys Ile Phe Gly Glu His Leu Gly Gly Thr Cys
      390         70                  75                  80
      392 ttc atc ctt ggt gaa cgc tac cca atc tgc tgc tac taa gcttgcagac   346
      393 Phe Ile Leu Gly Glu Arg Tyr Pro Ile Cys Cys Tyr
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 06/03/2002
PATENT APPLICATION: US/10/045,180A   TIME: 12:40:20

Input Set : A:\EP.txt
Output Set: N:\CRF3\06032002\J045180A.raw

<u>Please Note:</u>
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; N Pos. 85,143,670,970,1111,1150,1780,1974,2117,2133,2186,2191,2367
Seq#:1; N Pos. 4123

VERIFICATION SUMMARY   DATE: 06/03/2002
PATENT APPLICATION: US/10/045,180A   TIME: 12:40:20

Input Set : **A:\EP.txt**
Output Set: **N:\CRF3\06032002\J045180A.raw**

```
L:209 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:60
L:211 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:120
L:229 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:660
L:239 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:960
L:243 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:1080
L:245 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:1140
L:265 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:1740
L:271 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:1920
L:277 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:2100
L:279 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:2160
L:285 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:2340
L:343 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:4080
```